

## **Historic, Archive Document**

Do not assume content reflects current scientific knowledge, policies, or practices.



a 521  
R44A7

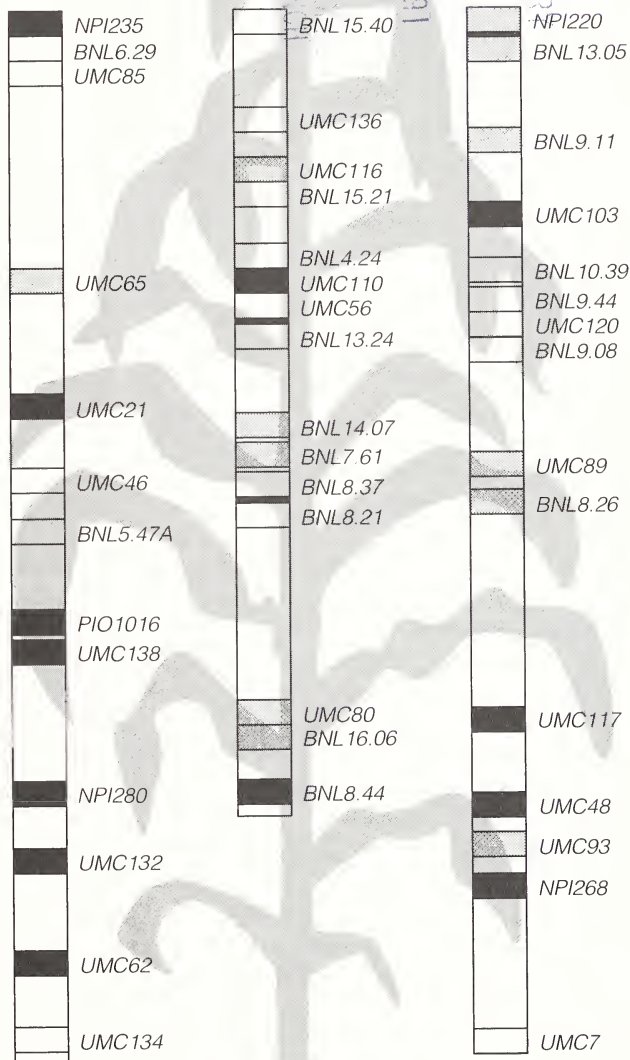
United States  
Department of Agriculture

Agricultural  
Research  
Service

ARS-94

June 1991

# USDA Plant Genome Research Program





# USDA Plant Genome Research Program

## Introduction

Plants power and drive our world. They convert light energy to products that directly or indirectly feed people and all other animals on Earth. They are the ultimate source of most of our energy resources, the fossil fuels oil, coal, and natural gas. They also supply nonfood commodities, such as wood products, fibers, medicines and pharmaceuticals, and industrial chemicals based on starches and oils. They cleanse the air of pollutants and capture gases that might otherwise contribute to a greenhouse effect. They provide the oxygen necessary for animal life.

Plant science applied to agriculture harnesses plants' potential by helping to ensure an adequate supply of agriculturally based food, fiber, industrial products, and clean air for the Earth's ever-growing population.

Genetics and other sciences have played a major role in agriculture over the years. Humans have improved cultivated plants by selection for thousands of years—more scientifically for the past 100—by applying the principles of heredity. The application of these genetic principles has enabled the rapid improvement of many crop varieties in the 20th century.

For many years, hybridization and selection techniques offered the chief means of crop improvement. Early progress produced substantial increases in yield, pest resistance, and other genetic traits of economic importance. But the giant steps are getting rarer, and traditional plant breeding technologies in some crops are reaching limits. Efficient methods to identify, isolate, and transfer desirable genes need to be developed to overcome these limitations of conventional breeding methods.

Since the late 1970's, the rapidly evolving tools of molecular biology have promised further improvement in crop and forest plants. A decade of genetic engineering research, however, has yet to yield the full potential of molecular biology in crop improvement.

Several bottlenecks have slowed biotechnological advancements in crop and forest species. Restricted basic knowledge of plant biochemistry, physiology, and pathology and lack of uniformly effective DNA transfer technologies and reproducible tissue culture procedures are barriers to advancement. Perhaps the most important of these is the difficulty in identifying and locating desired genes and gene systems in the plant genome. Better ability to locate and isolate specific genes will lead to more efficient breeding for important agronomic traits, such as drought and heat tolerance, pest resistance, reduced water contamination by pesticides and fertilizer, and increased quality and yield. Until those genes are located, precisely characterized, and transferred, systematic and predictable production of plants with useful characteristics will continue to be developed through trial- and- error research.

The technologies are already available to locate these genes. However, the job will require considerable financial and personnel resources. Plant gene mapping research is the key to effective use of these technologies and realization of biotechnology's potential.

In the testing of genetic principles, plants are unique. Short life cycles of most plant species and the presence of various nuclear genome configurations such as polyploidy or aneuploidy are distinct features. The abilities of many species to reproduce vegetatively and to self-fertilize are also different from what is typical of most animal species. Populations can be large enough that application of mathematical analysis offers greater sampling precision than is found in vertebrate animal systems. Chloroplasts and their inherent genetic systems also present special opportunities in plant genetic engineering and breeding programs.

Results from traditional breeding have also brought some serious consequences, such as resultant monocultures of successful cultivars that are especially susceptible to new diseases and pests. These and other factors could and have wiped out major plantings across whole regions. Luckily, breeding stock was available to counteract the losses in one major situation, the Southern corn blight in 1970. But time was lost breeding new disease-resistant lines into the population.

Explicit knowledge about genes and rapid exploitation of that knowledge can anticipate or neutralize diseases and problems with the interaction of harmful genotypes and the environment. With mapping information, new breeding schemes incorporating different specifically engineered gene complexes will be devised to help crop performance in each growing area. Plant genome research will also enable manipulation of genetic resources to achieve desirable variation. We will be able to capitalize more effectively on natural pest resistance in plants; even where this resistance is initially weak, genetic manipulation can make it more effective. The benefits of such capability would be lower costs to farmers and consumers and less chance of harming the environment with chemical pesticides.

The goal of the USDA Plant Genome Research Program is to find important genes and markers on chromosomes, isolate those genes from the genome, and transfer them rapidly to improve plant varieties. We have little knowledge of many important genes and their expressions. However, what we do know opens the way for a concerted genome research program. The Plant Genome Research Program will address single and multigenic traits that relate to agriculture, forestry, and environmental concerns (for example, global change and water quality). It will help maintain production stability and profitability and improve quality of agricultural products, maintain germplasm resources, and develop new crop plants and forest species. To accomplish these goals, the Plant Genome Research Program will foster and coordinate research. This

research will lead to the ability to identify, characterize, alter, and rapidly and precisely manipulate genes controlling traits of agricultural importance to meet society's needs.

Gene mapping will complement the U.S. National Plant Germplasm System, the National Biological Control Program, and other crop protection efforts involving genetic improvement. This goal recognizes agriculture as an industry based on growing plants for food, fiber, and industrial products. An additional outcome of the program will be the training of highly skilled U.S. scientists. The new scientists should be capable of using current and future discoveries in the biological sciences for the benefit of agriculture.

Release of products of agricultural biotechnology into the marketplace has been slower than expected. Apart from the fact that not enough time has elapsed for generation and release of new varieties, there are several reasons for the delay: Regeneration of most agronomically important species and transfer of genes in protoplasts and callus cultures remain significant challenges, though both are getting easier; gene transfer results are generally neither predictable nor stable; improving desirable traits requires finding the genes that code for the biochemical processes affecting those traits. Gene mapping and gene location—two very important activities needed to move biotechnology ahead—have only begun.

## **Scientific Objectives**

The program's objectives will be addressed by a competitive grant program managed cooperatively by the CSRS (USDA's Cooperative State Research Service) Competitive Grants Office and the ARS Office of Plant Genome Research Program.<sup>1</sup>

### **Objective 1—Mapping and Sequencing the Plant Genome**

Gene mapping research helps scientists use developments of molecular biology to move them into the next stage—product development. For example, gene mapping will make it possible to find and transfer genes that code for pathogen resistance, heat and drought stress, response to day length, and light intensity. Commodities with these improvements would reduce production costs, a benefit to farmers and consumers.

Plant genomes are different from the human/animal genome in several basic ways. This fact offers several challenges and opportunities to plant scientists. Two of the challenges:

---

<sup>1</sup>See Attachment A



- Chromosome separation and sorting are not always straightforward in plants because many species are polyploids.
- The chloroplast genome is unique to plants and thus presents an entirely new dimension to the study of genomes.

Novel opportunities arise primarily from the ability to breed plants in a manner that is not possible in animals. Chromosomal manipulation and cell fusion can produce hybrids from wide crosses. Plants can be selfed until the line becomes pure for certain traits, an impossible task in animals. Chromosomal changes or mutant lines can be developed that disrupt normal development and provide us with knowledge about plant growth.

The objectives are specified for initial and 5-year phases throughout this report.

## **Objective 1A—Construct genetic and physical maps**

Year 1—Begin mapping of crop and forest species.

Year 5—Complete maps of several crop species, at least at the 25-cM gap level.<sup>2</sup>

The state of genome mapping among agricultural and forest species ranges from preliminary to advanced. This section of the program supports the construction of maps for important crop and forest species. The goal is to construct maps of sufficient resolution that an adequate density of DNA probes is available to breeders as practical tools and to geneticists and genetic engineers as research aids.

To accomplish this phase of the mapping, an open competition for grants is the most effective vehicle. Which species to map is the investigator's decision.

Proposals will be judged by peer review. The primary concern is that the mapping will be done well and quickly. To be successful, a proposal must suggest ways to disseminate information and material to identified users. It should also describe methods of classifying and cataloging new information and materials. Proposals from multidisciplinary groups of molecular biologists, geneticists, and plant breeders who work together on a given species toward a common goal are desirable. A group proposal is likely to be viewed as having greater potential for

---

<sup>2</sup>cM=A unit of measurement for gene loci recombination frequencies



successful mapping research than submissions from individual scientists. Such multidisciplinary programs already exist for wheat, barley, soybeans, and cotton, among others.

## **Objective 1B—Detailed map or partial sequencing of specific regions of chromosomes**

Year 1—Solicit proposals and a profile research program dealing with specific chromosomal regions.

Year 5—Complete partial and total sequencing of important gene regions.

It is not necessary to wait for completion of a broad set of genomic maps of many crop and forest species before scientists can use modern technological advances for precise crop improvement. Therefore, the program will start using existing maps, DNA probes, and known gene sequences to identify and characterize specific simple and complex genes. Complex genes consist of many genes that work together to code for gene traits directly relevant to agriculture and forestry. Desirable genes to include are those that govern such traits as resistance and tolerance to diseases, pests, salt, and drought; hybrid vigor; allelopathy to weeds; and response to day length. Other important genes are those that control complex interactions between the environment and genotype and those that control recombination, the basis for plant breeding. As with the broader mapping program, successful completion of this phase requires multidisciplinary research. Plant biochemists, plant physiologists, entomologists, and plant pathologists must work together to identify the metabolic pathways involved in regulating expression of these genes.

Again, peer review will be instrumental in selecting successful proposals focused on specific genes and gene products.

## **Objective 1C—Development of technologies for mapping and sequencing**

Year 1—Initiate support of technology developments.

Year 5—Produce results on innovative genome research technologies.

Here proposals from individual investigators will have as high a chance of success as proposals from research groups.

Mapping and sequencing technologies are advancing rapidly, largely because of efforts by the National Institutes of Health/Department of Energy Human Genome Program. Plant genome researchers are building

on technological advances made in the Human Genome Program. Mapping and sequencing technology developments can meet the challenges and can take advantage of the uniqueness of plant systems. Examples of the new technologies that need development are:

- A method to identify sequences that will eliminate the need to store mapping probes
- Modifications and applications of the polymerase-chain-reaction (PCR)
- A new methodology to identify quantitative trait loci (QTL's)
- Computer software designed specifically for plant gene systems
- In-situ hybridization technologies for plant chromosomes
- Methods that allow mapping of polyploid genomes
- Chromosomal sorting and separation technologies specifically designed for complex plant genomes
- Methods to overcome obstacles to plant regeneration and transformation.

## **Objective 2—Databases and Biological Resource Management<sup>3</sup>**

Management of mapping/sequencing data and biological resources produced through genome research is a high priority for the Plant Genome Research Program. The volume of information expected from the research will be enormous, and the variety of biological data sets of species and related genetic resources of sequences, probes, and maps expected is extremely diverse. These results will be of value only if they are made readily and rapidly available to other researchers.

The clearest need is to establish a centralized and standardized database management system. Researchers must be able to easily deposit and retrieve mapping and sequencing data. An equally urgent need is to improve current genetic resource stock centers. The program will have to collect, store, and distribute new biological resources such as seeds, mutants, antibodies, proteins, and DNA probes. Establishment of new centers as materials accumulate is a major program challenge. An information clearinghouse is needed to maintain communication among all parties involved in the Plant Genome Research Program.

---

<sup>3</sup>See Attachment B, "Data Resources and the Plant Genome Research Program." If the report is not attached, please contact Dr. Keith Russell, The National Agricultural Library, Room 100, 10301 Baltimore, Blvd., Beltsville, MD 20705, (301) 344-3834.

## **Objective 2A—Stock center**

**Year 1**—Begin planning needed for specifying resource centers and what genetic resources are available and begin entering this information into a computer database.

**Year 5**—Have a complete database management system for access and dissemination.

As biological materials accumulate, they are likely to be different for the various species of agricultural and forest plants. Separate centers for major species groups are more manageable than a central stock center for all species. The Plant Genome Research Program will consider projects and contracts to organize such stock centers and associated databases and their management. The size of a stock center will vary depending on the need of each plant species. For crops such as maize, wheat, or tomatoes—for which many biological resources are already available—regional centralization of resource support is appropriate. For others crops, much more modest support in the form of supplemental funding to an individual researcher would be sufficient. A merit-based contract review will identify the best submissions regardless of the size. While a good center of this kind requires steady long-term support, each center supported by the Plant Genome Research Program will receive a thorough review by scientists every 3 to 5 years before renewal of support.

## **Objective 2B—Database management**

**Year 1**—Begin collecting plant genome data.

**Year 5**—Have a functional, debugged system for analysis, collection, and dissemination of plant genome data.

A major responsibility of the Plant Genome Research Program will be to facilitate general data analysis and management. Many of the recommendations of the Human Genome Joint Informatics Task Force (JITF) are relevant to plant genome research. The Plant Genome Research Program will have a group similar to the JITF, with some of the members serving on both committees. Needs common to both human and plant genome programs include:

- Methods for storage, retrieval, and analysis of nucleic acids and protein sequence
- High- and low-resolution physical maps

- Computer systems and storage methods (such as electronic notebooks) for quick and accurate handling of large amounts of experimental data generated in such projects
- Accessibility of the information in a variety of formats (tape, floppy disk, CD-ROM, and online) by persons with or without a high degree of computer expertise
- Support for automated sequencing and other large-scale mapping and sequencing efforts
- Development of links between the various databases and tools for comparison of data from different sources
- Standardization of data exchange and communication network protocols and integration of database schemes

Needs exclusive to plant genome research include:

- Information about genotype of many genetic strains of a wide variety of organisms and their availability from germplasm banks and organism-specific genetic stock centers.
- Ability to support genetic descriptions (genotype, phenotype, and standardization of gene nomenclature) of diverse organisms with differing genetic systems and different experimental histories, reflecting approaches to different genetic questions. A plant species may be monoecious or dioecious; apomictic, self-fertile, or outcrossing; amenable to single-cell cloning and regeneration; haploid, diploid, triploid, or polyploid in some or all tissues; an angiosperm or a gymnosperm. It may have a life cycle of weeks or centuries. Amounts of information will vary on cytogenetic, molecular, Mendelian, and practical breeding and on its physiological-genetic, population, ecological, or evolutionary genetics. This information, assessing the user's needs for access to such information, and developing mechanisms for appropriate access are major challenges for the plant genome program.
- Cross-species comparisons of agriculturally important genes—for example genes affecting resistance to drought, disease, and pests; symbiotic interactions; and storage protein or carbohydrate biosynthesis places complex demands on data management efforts.

A pilot effort involving three plant species (two crops and a forest species) will be very useful for assessing the existing or needed databases for each species, the databases and software needed to link information from the three species, and problems of database structures and management. The pilot effort will be supported for the first 3 years only, and then it will be evaluated for application to the genome program.

There is a need to establish interdisciplinary programs for improving communication between computer scientists and plant biologists. Better communication would provide a pool of scientists with appreciation of both areas as well as expertise in one or both, who are capable and willing to undertake the interactions required to develop the needed systems. The Plant Genome Research Program will encourage and support such interdisciplinary activities.

## **Objective 2C—Placing products in the marketplace in 5 to 10 years**

Year 1—Institute a policy among government, industry, and universities for technology transfer.

Year 5—Have operational system to transfer knowledge about plant genomes to commercialize in the marketplace.

As the gene mapping and sequencing technologies and the actual maps are developed, it should be possible to apply them to identify, locate, and isolate genes for specific traits of interest and to use them to manipulate crop plants by genetic engineering or conventional breeding. Currently, many areas in plant biology lack basic information on plant metabolism and structure, and this must be rectified before the mapping/sequencing technology and information can be used to their full potential. Ongoing research at universities, ARS, and the USDA Competitive Research Grants Program can fill the gaps in basic knowledge. In 5 years, the state of science will likely allow us to apply mapping information for genetic engineering and plant breeding on a large scale. Because there is no major or widespread ethical, moral, or social rejection of genetically engineered plants, we expect the technology transfer to succeed in the marketplace.

With the knowledge gained through this program, we will move closer to being able to design crops that meet market needs and that are increasingly resistant to the vagaries of weather and the degradation of pests. The eventual result will be a market-driven agriculture.

